

Novel Kazal-type Protease inhibitors from the skin secretion of the Splendid leaf frog, *Cruziohyla calcarifer*

Proaño Bolaños, C., Li, R., Zhou, M., Wang, L., Xi, X., Tapia, E. E., Coloma, L. A., Chen, T., & Shaw, C. (2017). Novel Kazal-type Protease inhibitors from the skin secretion of the Splendid leaf frog, *Cruziohyla calcarifer*. *EuPA Open Proteomics*, 15, 1-13. <https://doi.org/10.1016/j.euprot.2017.02.001>

Published in:
EuPA Open Proteomics

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2017 the authors.

This is an open access article published under a Creative Commons Attribution-NonCommercial-NoDerivs License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits distribution and reproduction for non-commercial purposes, provided the author and source are cited.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.



Novel Kazal-type proteinase inhibitors from the skin secretion of the Splendid leaf frog, *Cruziohyla calcarifer*



Carolina Proaño-Bolaños^{a,c,*}, Renjie Li^a, Mei Zhou^a, Lei Wang^a, Xinping Xi^a,
Elicio E. Tapia^b, Luis A. Coloma^{b,c}, Tianbao Chen^a, Chris Shaw^{a,**}

^a Natural Drug Discovery Group, School of Pharmacy, Queen's University Belfast, 97 Lisburn Road, BT9 7BL, Belfast, Northern Ireland, UK

^b Centro Jambatu de Investigación y Conservación de Anfibios, Fundación Otonga, Geovanni Farina 566 y Baltra, San Rafael, Quito, Ecuador

^c Universidad Regional Amazónica Ikiam, Muyuna, Tena, Ecuador

ARTICLE INFO

Article history:

Received 24 April 2016

Received in revised form 4 January 2017

Accepted 15 February 2017

Available online 27 February 2017

Keywords:

Kazal-type proteinase inhibitors

Frog skin secretion

Peptidomic

Molecular cloning

Tandem mass spectrometry

Cruziohyla calcarifer

ABSTRACT

Peptidase inhibitors have an important role controlling a variety of biological processes. Here, we employed a peptidomic approach including molecular cloning, tandem mass spectrometry and enzymatic assays to reveal 7 Kazal-type proteinase inhibitors (CKKPs) (18 variants) in the skin secretion of the unexplored frog, *Cruziohyla calcarifer*. All 18 proteins shared the Kazal pattern C-X₍₇₎-C-X_(6,7)-C-X_(6,7)-Y-X₍₃₎-C-X₍₂₎-C-X₍₁₅₋₂₁₎-C and 3 disulphide bridges. Based on structural comparative analysis, we deemed trypsin and chymotrypsin inhibitory activity in CKKP-1, 4 and CKKP 2, 5, 7, respectively. These peptidase inhibitors presumably play a role to control the balance between other functional peptides produced in the amphibian skin secretions.

© 2017 The Authors. Published by Elsevier B.V. on behalf of European Proteomics Association (EuPA). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Amphibian skin contains granular glands that are responsible for synthesizing a complex mixture of different biologically active peptides. Among these peptides are tachykinins, bradykinins, sauvagines, caeruleins, bombesins, dermorphins, dermaseptins, and other peptides [1]. These peptides are synthesized as inactive precursors often containing a signal peptide, an acidic spacer, and the mature sequence. In addition, they usually have one or more predicted processing sites within the acidic spacer, such as lysine-arginine (KR) or two arginines (RR), which are a target for peptidases with trypsin-like specificity [2–4].

Peptidases are hydrolases that are crucial in controlling a variety of biological process, such as digestion, blood clotting, wound healing, and host defence; but they are also involved in pathogenic infection, viral replication, and disease progression [5–7]. Besides, proteolytic enzymes play an important role in

releasing active mature peptides upon the surface of the frog, and for this reason these enzymes must be carefully controlled. Among the processes that could be involved in their control are spatial and temporal regulation, zymogen activation, peptidase degradation, and macromolecular peptidase inhibitors. However, these mechanisms have not been proven in amphibians up to now [5–12].

Peptidase inhibitors as well as peptidases are ubiquitous in nature. Due to the fact that peptidase inhibitors show a high level of homology in their active sites, they can inhibit multiple peptidases at a rate of 1:5. The mechanism of inhibition in the majority of peptidase inhibitors is competitive, interacting with the active site of the peptidase in a substrate-like manner [6].

Proteinase inhibitors of different families and specificities have been found in some amphibian species. The first proteinase inhibitors were found in the skin secretions of five species of bombinid toads, including the following trypsin/thrombin inhibitors: BSTI of *Bombina bombina*; BMS11 and BMS12 of *B. microdeladigitata*; BOTI of *B. orientalis*; BVTI of *B. variegata*; and BMTI of *B. maxima*. They have 60 amino acids and masses between 6345 and 6446.56 Da. Noticeably, they all contain 10 cysteine residues similar to the distribution pattern of the peptidase inhibitor of the nematode worm, *Ascaris suum*. In addition, they share the reactive centre motif -CDKK-, and their K_i is in the range of 0.1–1 μM

* Corresponding author at: School of Pharmacy, Queen's University Belfast, 97 Lisburn Road, BT9 7BL, Belfast, Northern Ireland, UK.

** Corresponding author.

E-mail addresses: cproanobolanos01@qub.ac.uk, carolina.proano@ikiam.edu.ec (C. Proaño-Bolaños), chris.shaw@qub.ac.uk (C. Shaw).

[8,9,13,14]. Another species, *Rana areolata* contains three peptidase inhibitors: one trypsin inhibitor of 61 amino acids that also contains the 10 cysteine residue motif (6751 Da) and is similar to the protein inhibitor from *A. suum*; a 48 residue elafin-related peptide (of mass 5164 Da) with 8 cysteine residues distinctive for whey acidic protein (WAP) motif; and a secretory leukocyte peptidase inhibitor [15].

Other proteinase inhibitors found in amphibian skin secretions can be classified based on the presence of their structural motifs such as Kunitz, Bowman-Birk and Kazal types. Two Kunitz type proteinase inhibitors have been isolated from *Dyscophus guineti* and *Kassina senegalensis*. The first is a trypsin inhibitor of 57 residues and 6301 Da, which belongs to the Kunitz/bovine pancreatic trypsin inhibitor family; and the second a chymotrypsin inhibitor (KSCI) of 62 residues and 6776.24 Da, which is structurally similar to other chymotrypsin inhibitors from silkworms to snakes. KSCI has its P1 site occupied by a phenylalanine residue and contains 6 cysteines that form 3 disulphide bridges [16,17].

The next group of proteinase inhibitors contain the Bowman-Birk motif. They include the following trypsin inhibitors: HV-BBI isolated from *Odorrana versabilis*; HJTI isolated from *O. hejianensis*; and pLR-HL isolated from *Hylarana latouchii*. These inhibitors have 17–18 residues, their masses range between 1804.83 and 2013.95 Da, and their K_i varies between 19 and 388 nM. Besides, they share a disulphide loop between 2 cysteines spaced by 9 residues, and have a lysine in the third position inside the loop, which corresponds to the P1 site of the inhibitor [18,19]. In addition, OGTI isolated from *O. grahami*, is another trypsin inhibitor similar to those described above. OGTI contains 17 amino acids and its molecular mass is 1949.4 Da. However, OGTI contains a smaller disulphide loop formed by 2 cysteines that are 4 residues apart, and it has a lysine in the P1 site as the immediate residue following the first cysteine [20].

Finally, Kazal-type protein inhibitors seem to be more taxo-specific. Until now, they have only been reported in the Phyllomedusinae clade. Two prolyl endopeptidase inhibitors, PSKP-1 and PSKP-2 of 58 residues and 6695.87 and 6548.65 Da, respectively, have been isolated from *Phyllomedusa sauvagii*. PSKP-1 and PSKP-2 share two prolines in the P1 and P2 positions of the putative active sites, making them unable to inhibit trypsin, chymotrypsin, v8 peptidase, and proteinase K [21]. In contrast, ACKTI of 5892.82 Da – a trypsin inhibitor isolated from *Agalychnis callidryas* – has proline in P2 but arginine in the P1 position consistent with other trypsin inhibitors [22]. In addition, another two Kazal-type peptidase inhibitors – PI01 and PI02 – were identified in *P. nordestina* by ETS analysis; however, their specificity has not been elucidated [23].

Although the biological roles of the amphibian proteinase inhibitors have not been established with certainty, they may include defence against extrinsic peptidases produced by pathogenic microorganisms to prevent damage of host tissue and evasion of host defences. Proteinase inhibitors could also prevent degradation of bioactive peptides, so they can target cell receptors. In addition, proteinase inhibitors might act indirectly as regulators of the processing reactions of bioactive peptides, including cationic α -helical antimicrobial peptides, allowing them to be released onto the skin, so they can display their activity and protect the skin from invading microorganisms [8,9,16].

The present study was focused on the Splendid leaf frog, *Cruziohyala calcarifer*, which belongs to the Phyllomedusinae clade – a known source of pharmacological and antimicrobial peptides. Recently, one insulin-releasing peptide RK-13, and 18 cruzioseptins with antimicrobial activity have been described from *C. calcarifer* [24,25]. Here, we describe a group of 7 peptidase inhibitors (with 18 variants) in *C. calcarifer* skin secretion which belong to the Kazal-type family. One of these, CCKP-1, showed trypsin inhibitory

activity and possessed a lysine in its P1 site and unusually, an asparagine in its P2 site. Based on their structural homology, it is predicted that CCKP-2, CCKP-5 and CCKP-7 have chymotrypsin inhibitory activity, while CCKP-4 has trypsin inhibitory activity. Therefore, the proteinase inhibitors of Kazal-type from *C. calcarifer* are the most diverse group of proteinase inhibitors found to date in a single amphibian species.

2. Material and methods

2.1. Sourcing of samples

The skin secretions of *Cruziohyala calcarifer* employed in this study came from two different geographical locations, Costa Rica and Ecuador.

The first Costa Rican sample consisted of a pool of two adults secretions collected in 1999. While the Ecuadorian sample consisted of a pool of four juvenile captive breed and one wild adult secretion collected in 2013. The four juvenile captive bred frogs ($n=4$) (from Esmeraldas Province, Reserve Otokiki) were provided by Centro Jambatu for Research and Conservation of Amphibians in Ecuador and the wild specimen of *Cruziohyala calcarifer* ($n=1$) was collected in northwestern Ecuador (Esmeraldas Province, Durango). Skin secretions were extracted from each frog by lightly stressing the animal – massaging the dorsal area of the frog – then washing off the secretion with distilled water. Secretions from the same geographical location were pooled, equally split into two 50 mL conical tubes, and then freeze dried. Samples were transported to Queen's University Belfast at room temperature and dried samples were stored at -20°C prior to their analysis.

Moreover, twelve additional samples were extracted from a group of 13-month-old captive bred frogs in 2015. The parental line of these frogs came from a Costa Rican population and the animals were housed in terraria as pets in Belgium and Austria. Skin secretions were extracted in the same way as described above but kept individually. In contrast to the previous samples, the twelve samples were acidified with TFA and transported at room temperature to the laboratory facilities in Queen's University Belfast where they were freeze-dried.

Collection and rearing of frogs in Ecuador and transportation of samples were done under permits of the Ecuadorian Ministerio de Ambiente (MAE) (described in acknowledgments).

2.2. Molecular cloning

One aliquot containing half of the dried secretion material of the Ecuadorian sample was dissolved in 1 mL of cell lysis/binding buffer and polyadenylated mRNA was isolated using magnetic Dynabeads Oligo (dT)s, as described by the manufacturer (Dynal Biotec, UK). Isolated mRNA was subjected to 3'-rapid amplification of cDNA by using the SMART-RACE kit (Clontech, UK). Briefly, the 3'-RACE reaction used a nested universal primer (NUP), provided with the kit, and two senses primers: S1: 5'-AGCAGCAAAGAAGAAGGCCATG-3' and S2: 5'-GAGAAGAAGCCATGAAGACTCTGA-3', that were complementary to the signal sequence of the ACKTI gene precursor of *Agalychnis callidryas*. The 3'-RACE product was purified and cloned using a pGEM-T vector system (Promega Corporation), and later sequenced using an ABI 3100 automated sequencer. Nucleotide sequences were analysed with MEGA 6.0 and Vector NTI software (Life technologies) [26]. Sequences were compared with databases in the NCBI using the BLAST tool [27]. Theoretical peptide masses were calculated with the peptide mass calculator 3.2 and secondary structure was predicted with GOR IV method [28,29]. Nucleotide sequences were submitted to the GenBank of the NCBI, accession numbers are included in Table 1.

Table 1
Primary structures of 18 variants of *Kazal* type proteins from *Cruziolihyla calcarifer* as confirmed by tandem mass spectrometry. Data represent the best scores of 13 repetitions.

Peptide	Sequence	Theoretical Average MW (Da)	LCQ MW (Da)	# Peptide fragments	#AAs	Coverage %	Score	Accession number
CCRP-1	AVSAECARYGLACNKLAPVCGTGDGTTYSNQCMCLCYNNRKNKNIERSRGR	5920.82	5922.79	1631	53	100	17.80	KX065060
CCRP-2	ATEPDCCKYKPLAQNPVCGTDRMYNECALCVPMRDSKNVKKIJKMGKC	6198.33	6199.99	902	55	100	8.68	KX065061
CCRP-3	ATPKPCLFSSGCPSTQDFVCGTDSNMYNECALCVPMRDSKNVKKIJKMGKC	5780.69	5782.63	2335	54	100	42.37	KX065062
CCRP-4a	GGVLLDCRPVPCSKIFDPVCGTFTYDNTCELCCKAQRENPRISMRTKGC	5992.94	5994.90	699	54	100	7.52	KX065063
CCRP-4b	VRLDCRPVPCSKIFDPVCGTFTYDNTCELCCKAQRENPRISMRTKGC	5921.87	5923.87	647	52	100	8.90	KX065064
CCRP-4c	VRLDCRPVPCSKIFDPVCGTFTYDNTCELCCKAQRENPRISMRTKGC	7446.60	7447.70	864	64	100	20.52	KX065065
CCRP-5a	VIENCKYEGKKDLNPNVCGTNGREYNECALCVFIRDSKKADKMKIKKMGKC	6667.84	6669.29	649	58	100	12.57	KX065066
CCRP-5b	VIENCKYEGKKDLNPNVCGTNGREYNECALCVFIRDSKKADKMKIKKMGKC	6668.87	6670.31	480	58	100	14.77	KX065067
CCRP-5c	VIENCKYEGKKDLNPNVCGTNGREYNECALCVFIRDSKKADKMKIKKMGKC	6635.82	6637.34	684	58	100	11.40	KX065068
CCRP-5d	VIENCKYEGKKDLNPNVCGTNGREYNECALCVFIRDSKKADKMKIKKMGKC	6636.76	6638.30	681	58	100	10.53	KX065069
CCRP-6a	EEDVCPWYVFGCHDKYTVCGTDCGYPNKCTCKINGEDNIIKIKWGC	5849.58	5851.58	426	51	100	8.46	KX065070
CCRP-6b	EEDVCPWYVFGCHDKYTVCGTDCGYPNKCTCKINGEDNIIKIKWGC	5875.59	5877.65	617	51	100	7.86	KX065071
CCRP-7a	PLPSQPFKKVLTALAEPNCKYEGKKDLNPNVCGTNGRTYNECALCVFIRDSKKADKMKIKKMGKC	8377.81	8378.27	865	73	100	13.16	KX065072
CCRP-7b	PLPSQPFKKVLTALAEPNCKYEGKKDLNPNVCGTNGRTYNECALCVFIRDSKKADKMKIKKMGKC	8272.54	8273.11	1015	73	100	18.61	KX065073
CCRP-7c	PLPSQPFKKVLTALAEPNCKYEGKKDLNPNVCGTNGRTYNECALCVFIRDSKKADKMKIKKMGKC	8376.83	8377.29	1001	73	100	13.46	KX065074
CCRP-7d	PLPSQPFKKVLTALAEPNCKYEGKKDLNPNVCGTNGRTYNECALCVFIRDSKKADKMKIKKMGKC	8391.84	8392.29	868	73	100	17.68	KX065075
CCRP-7e	PLPSQPFKKVLTALAEPNCKYEGKKDLNPNVCGTNGRTYNECALCVFIRDSKKADKMKIKKMGKC	8278.68	8279.19	1113	73	100	14.51	KX065076
CCRP-7f	PLPSQPFKKVLTALAEPNCKYEGKKDLNPNVCGTNGRTYNECALCVFIRDSKKADKMKIKKMGKC	8316.64	8317.18	961	73	100	21.25	KX065077

2.3. Reverse phase HPLC fractionation of skin secretion

Different experimental conditions were necessary for the separations as follow:

The Ecuadorian sample (remaining half of the dried secretion) was dissolved in 1.2 mL of solution A (99.95% H₂O, 0.05% trifluoroacetic acid) and clarified by centrifugation. 1 mL supernatant was subjected to reverse phase HPLC employing a Waters binary pump HPLC system fitted with an analytical column: Phenomenex Jupiter C-18 (4.6 × 250 mm). Peptides were eluted with a gradient formed from 100% solution A (99.95% H₂O, 0.05% trifluoroacetic acid) to 100% solution B (80.00% acetonitrile, 19.95% H₂O, 0.05% trifluoroacetic acid) for 240 min at a flow rate of 1 mL/min. Fractions of 1 mL were collected every minute. Detection at 214 and 280 nm was performed continuously.

The Costa Rican sample (0.5 mL of 3 mL) from 1999 was injected onto an HPLC system fitted with a semi-preparative C-18 Phenomenex Jupiter Column employing a gradient of 0–80% solution B for 80 min at a flow rate of 2 mL/min. Fractions were collected every minute. This analysis was performed on 08/03/1999.

2.4. MALDI-TOF MS

The molecular masses of peptides and proteins in each chromatographic fraction were analysed by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) using a linear time of flight Voyager DE mass spectrometer (Perceptive Biosystems, MA, USA). First, the instrument was set in positive detection mode and all fractions were analysed with α -cyano-4-hydroxycinnamic acid (CHCA) matrix (10 mg/mL) for low mass peptides. The settings of the instrument were as follows: mode of operation: linear, accelerating voltage: 2000 V, acquisition mass range: 500–5000 Da, number of laser shots: 50/spectrum, laser intensity: 1813. Later, fractions were analysed for high mass proteins employing sinapinic acid (10 mg/mL) as matrix. The settings of the instrument were as follow: mode of operation: linear, accelerating voltage: 25000 V, acquisition mass 500–10,000 Da, number of laser shots: 50/spectrum, laser intensity: 3119. In both cases, two microliters of sample plus 1 μ L of matrix (10 mg/mL) were allowed to dry and later analysed for mass/charge.

2.5. MS/MS sequencing

20 μ L of the remaining skin secretion of the Ecuadorian sample dissolved in solution A were applied directly to an analytical HPLC column (Phenomenex C-18; 4.6 × 150 mm) connected to a linear ion-trap mass spectrometer (LCQ-Fleet, Thermo Fisher). The linear elution gradient was formed from 100% solution A (99.90% H₂O, 0.1% formic acid) to 100% solution B (80% acetonitrile, 19.0% H₂O, 0.1% formic acid) for 135 min at a flow rate 20 μ L/min. The mass spectrometer was controlled by Xcalibur software (Thermo). The mass analysis was performed in positive ion mode, and the acquired spectra were in the range of m/z 500–2000 Da. The parameters for electrospray ionization ion-trap mass spectrometry (ESI/MS) were: spray voltage +4.5 kV, drying gas temperature 320 °C, drying gas flow 200 μ L/min, and maximum accumulation time – for the ion trap – 350 ms. After the first mass analysis in full scan mode, peptide ions with >50% relative intensity were fragmented by collision induced dissociation (CID), in order to generate b and y ions that were detected in a second mass analysis. Fragment ion profiles were analysed by Proteome Discoverer 1.0 software (Thermo). In order to confirm the amino acid sequences of individual peptides we used the Sequest™ algorithm which compares the acquired fragment ion profiles with the theoretical

fragment ions generated from a FASTA file containing the specific sequences obtained by molecular cloning of this species.

In addition, twelve individual samples from the Costa Rican population were analysed in the same way as described above but instead of 135 min these samples were run for only 120 min.

2.6. In solution Trypsin digestion

The HPLC fraction 91 was freeze-dried and later reconstituted in 5 µL of Solution A (99.95% H₂O; 0.05% TFA). The sample was denatured with 8M Urea, 25mM Tris and 20mM DTT (final concentrations) and incubated at 60°C for 1 h. The alkylation of the sample was performed with Iodoacetamide (IAA) at 40mM and

A) CCKP-1

	M	K	T	L	I	S	S	A	L	L	F	C	I	L
1	GAAGAAGCCA	TGAAGACTCT	GATCTCCTCC	GCCTTGCTCT	TCTGTATACT									
	<u>A A A</u>	<u>L W P V</u>	<u>P E A</u>	<u>E G A</u>	<u>V S A E</u>									
51	GGCAGCAGCA	TTGTGGCCGG	TTCCTGAAGC	AGAAGGTGCA	GTTTCGGCAG									
	<u>C A R</u>	<u>Y G L</u>	<u>A C N K</u>	<u>M L A</u>	<u>P V C</u>									
101	AATGCGCCCG	CTATGGTCTG	GCCTGCAACA	AAATGTTAGC	ACCTGTTTGT									
	<u>G T D G</u>	<u>T T Y</u>	<u>S N Q</u>	<u>C M L C</u>	<u>Y Y N</u>									
151	GGAAGTATG	GAACGACATA	CTCCAACCAG	TGCATGCTGT	GCTACTATAA									
	<u>R K N</u>	<u>K K N I</u>	<u>E I R</u>	<u>S R G</u>	<u>R C</u>	*								
201	CAGAAAAAAT	AAGAAGAATA	TTGAAATTAG	GTCAAGGGGT	CGATGCTGAG									
251	GAAGTCGTAT	TAGATCCCGG	AGAGCTTCAG	ACACCGATGT	TCTGTGAAGT									
301	GATGATTACG	GCAAAAGAAA	AATCAATGAC	ATGAAAAATA	AATCCTGTGC									
351	AATCCAATGT	CCAGACATGT	TACAATAAAG	AGTGGAAGCC	TTATAAAAAA									
401	AAAAAAAAAA	AAAAAAAAAA												

B) CCKP-2

	M	K	T	L	I	S	S	A	L	L	F	C	I	L	A	A
1	GCCATGAAGA	CTCTGATCTC	CTCCGCCTTG	CTCTTCTGTA	TACTGGCAGC											
	<u>A L W</u>	<u>P V P E</u>	<u>A E G</u>	<u>A T E</u>	<u>P D C K</u>											
51	AGCATTGTGG	CCAGTTCCTG	AAGCAGAAGG	TGCAACTGAG	CCAGACTGTA											
	<u>K Y P</u>	<u>G K C</u>	<u>P L A Q</u>	<u>N P V</u>	<u>C G T</u>											
101	AGAAGTATCC	GGGGAATGT	CCTTTAGCTC	AAAATCCTGT	GTGTGGAACA											
	<u>D G R M</u>	<u>Y Y N</u>	<u>E C A</u>	<u>L C V F</u>	<u>M R D</u>											
151	GATGGAAGAA	TGTACTACAA	TGAGTGCGCG	CTGTGCGTCT	TTATGAGGGA											
	<u>S K N</u>	<u>K V K I</u>	<u>Q I K</u>	<u>K M G</u>	<u>K C</u>	*										
201	CAGTAAGAAT	AAAGTTAAAA	TTCAAATTAA	AAAAATGGGT	AAATGCTGAA											
251	GAGCTCGTAT	TAAATCCTGG	AGAGCTTCAG	ACACCGATGT	TCTGCAAGT											
301	GATGATAAAA	ACAGCGAATA	AAAAATCAAT	GACATGAAAA	ATAAATCCTG											
351	TGCAATCCAA	TGTCCAGACA	TGTTACAATA	AAGAGTGGA	GCCTTGATCT											
401	GTAGCAACAA	AAAAAAAAAA	AAAAAAAAAA	AAA												

C) CCKP-3

	M	K	T	L	I	S	S	V	L	L	F	C	I	L	A	A	A
1	ATGAAGACTC	TGATCTCCTC	CGTCTTGCTC	TTCTGTATAC	TGGCAGCAGC												
	<u>L W P</u>	<u>V P K A</u>	<u>E G A</u>	<u>T K P</u>	<u>K C P S</u>												
51	ATTGTGGCCA	GTTTCTAAAG	CAGAAGGTGC	AACTAAGCCG	AAATGTCCCT												
	<u>L F S</u>	<u>S G C</u>	<u>P S T Q</u>	<u>D F V</u>	<u>C G T</u>												
101	CCCTGTTTTC	CTCAGGGTGC	CCCAGTACAC	AAGATTTTGT	TTGTGGAACA												
	<u>D G N S</u>	<u>Y M N</u>	<u>E C V</u>	<u>M C K M</u>	<u>N K N</u>												
151	GATGGAATTT	CATACATGAA	TGAGTGCGTG	ATGTGCAAAA	TGAACAAGAA												
	<u>N G G</u>	<u>K V K V</u>	<u>V K K</u>	<u>G Y C</u>	*												
201	CAATGGGGGA	AAGGTTAAGG	TTGTAAAAAA	GGGATACTGC	TGAGAAGCTC												
251	ATATTAAATC	CTGGAGACAC	CGATGTTCTG	TGAAGTGATG	ATTACAGCAA												
301	AAGGAGAATT	AATCAAATGA	AAATTCCTGT	GCAATCCAAT	GTCCAGACAT												
351	GTTACAATAG	AGAGTGGAAG	CCTTGATCTG	TAACAAAAAA	AAAAAAAAAA												
401	AAAAAAAAAA	A															

Fig. 1. Nucleotide and translated open-reading frame amino acid sequences of cloned cDNAs that encode the biosynthetic precursors of the Kazal-type proteins from *Cruziophyla calcarifer*. A–G) Representative seven types of Kazal-type proteins with a range of 51–73 amino acids. The putative signal peptides are double-underlined, the mature sequences are single-underlined and the stop codons are indicated by asterisks.

D) CCKP-4a

		<u>M</u>	<u>K</u>	<u>T</u>	<u>L</u>	<u>I</u>	<u>S</u>	<u>S</u>	<u>A</u>	<u>L</u>	<u>L</u>	<u>F</u>	<u>C</u>	<u>I</u>	<u>L</u>	<u>A</u>
1	GAAGCCATGA	AGACTCTGAT	CTCCTCCGCC	TTGCTCTTCT	GTATACTGGC											
	<u>A</u>	<u>A</u>	<u>L</u>	<u>L</u>	<u>P</u>	<u>V</u>	<u>L</u>	<u>E</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>V</u>	<u>V</u>	<u>L</u>	<u>L</u>	<u>D</u>
51	AGCAGCATTG	TTGCCGGTTC	TTGAAGCAGG	AGGTGTAGTT	CTGCTGGACT											
	<u>R</u>	<u>P</u>	<u>Y</u>	<u>G</u>	<u>P</u>	<u>V</u>	<u>C</u>	<u>S</u>	<u>K</u>	<u>I</u>	<u>F</u>	<u>D</u>	<u>P</u>	<u>V</u>	<u>C</u>	<u>G</u>
101	GCCGCCCGTA	TGGCCCAGTT	TGCAGCAAAA	TATTTGATCC	TGTTTGTGGA											
	<u>T</u>	<u>N</u>	<u>F</u>	<u>I</u>	<u>T</u>	<u>Y</u>	<u>D</u>	<u>N</u>	<u>T</u>	<u>C</u>	<u>E</u>	<u>L</u>	<u>C</u>	<u>K</u>	<u>A</u>	<u>Q</u>
151	ACAAATTTCA	TAACATATGA	CAATACGTGC	GAGCTGTGCA	AGGCACAAAG											
	<u>E</u>	<u>N</u>	<u>P</u>	<u>R</u>	<u>I</u>	<u>S</u>	<u>M</u>	<u>R</u>	<u>T</u>	<u>K</u>	<u>G</u>	<u>K</u>	<u>C</u>	<u>*</u>		
201	AGAGAATCCG	AGGATTAGTA	TGAGAACTAA	AGGTAAATGC	TGATGGAAGT											
251	TGTATCAGAT	CCCGGAGAGC	TTCAGACGCT	GACTTTTCTT	TTGCCAAAGT											
301	CATCGTTTCA	GTTACATGTA	AAATAAACCC	TGTGCAATCC	AATGTCCAGA											
351	CATGTTCCAA	TAAAGAGTGG	AAGCCTTGAT	CTGCAAAAAA	AAAAAAAAAA											
401	AAAAAAA															

E) CCKP-5a

		<u>M</u>	<u>K</u>	<u>T</u>	<u>L</u>	<u>I</u>	<u>S</u>	<u>S</u>	<u>V</u>	<u>L</u>	<u>L</u>
1	GCAGCAAAAG	AAGAAGAAGC	CATGAAGACT	CTGATCTCCT	CCGTCTTGCT						
	<u>F</u>	<u>C</u>	<u>I</u>	<u>L</u>	<u>A</u>	<u>T</u>	<u>A</u>	<u>L</u>	<u>W</u>	<u>P</u>	<u>I</u>
51	CTTCTGTATA	CTGGCAACAG	CATTGTGGCC	AATTCATGAA	GCAGAAGCTG						
	<u>I</u>	<u>E</u>	<u>P</u>	<u>N</u>	<u>C</u>	<u>K</u>	<u>K</u>	<u>Y</u>	<u>E</u>	<u>G</u>	<u>K</u>
101	TAATTGAGCC	AACTGTAAAG	AAGTATGAGG	GAAAAAGTG	TGATCTAAAC						
	<u>P</u>	<u>N</u>	<u>P</u>	<u>V</u>	<u>C</u>	<u>G</u>	<u>T</u>	<u>N</u>	<u>G</u>	<u>R</u>	<u>E</u>
151	CCAAATCCTG	TTTGTGGAAC	AAATGGACGA	GAATACTTCA	ATGAGTGCGC						
	<u>L</u>	<u>C</u>	<u>V</u>	<u>F</u>	<u>I</u>	<u>R</u>	<u>D</u>	<u>S</u>	<u>K</u>	<u>K</u>	<u>K</u>
201	GCTGTGCGTC	TTTATCAGGG	ACAGTAAGAA	GAAAGCTGAT	AAGATGGTTA						
	<u>I</u>	<u>K</u>	<u>K</u>	<u>W</u>	<u>G</u>	<u>K</u>	<u>C</u>	<u>*</u>			
251	AAATTAATAA	ATGGGGTAAA	TGCTGAGGAG	CTCATAATAA	ATCCTGGAGA						
301	GCTTCAGACA	CCGATGTTCT	GTGAAGTGAT	GATTACAGCA	AAAGAAAAAT						
351	CAATCAGATG	AAAAATAAAT	CCTGTGCAAT	CCAATGTCCA	GACATGTTAC						
401	AATAAAGAGT	GGAAGCCTTA	AAGAAAAAAA	AAAAAAAAAA	AAAAAAAAAA						

F) CCKP-6a

		<u>M</u>	<u>K</u>	<u>T</u>	<u>L</u>	<u>I</u>	<u>S</u>	<u>S</u>	<u>A</u>	<u>L</u>	<u>L</u>	<u>F</u>	<u>C</u>	<u>I</u>	<u>L</u>	<u>A</u>
1	GAAGCCATGA	AGACTCTGAT	CTCCTCCGCC	TTGCTCTTCT	GTATACTGGC											
	<u>A</u>	<u>A</u>	<u>L</u>	<u>W</u>	<u>P</u>	<u>I</u>	<u>P</u>	<u>E</u>	<u>V</u>	<u>E</u>	<u>G</u>	<u>E</u>	<u>E</u>	<u>D</u>	<u>V</u>	<u>A</u>
51	AGCAGCATTG	TGGCCAATTC	CTGAAGTAGA	AGGTGAAGAA	GATGTGGCCT											
	<u>P</u>	<u>W</u>	<u>Y</u>	<u>Y</u>	<u>V</u>	<u>F</u>	<u>G</u>	<u>C</u>	<u>H</u>	<u>D</u>	<u>K</u>	<u>Y</u>	<u>T</u>	<u>V</u>	<u>C</u>	<u>G</u>
101	GTCCTTGGA	TTACGTGTTT	GGGTGCCACG	ACAAATACAC	TGTTTGTGGA											
	<u>T</u>	<u>D</u>	<u>G</u>	<u>V</u>	<u>T</u>	<u>Y</u>	<u>P</u>	<u>N</u>	<u>K</u>	<u>C</u>	<u>T</u>	<u>L</u>	<u>C</u>	<u>K</u>	<u>I</u>	<u>N</u>
151	ACAGATGGAG	TCACATACCC	CAATAAATGC	ACCCTGTGCA	AGATAAATGG											
	<u>E</u>	<u>D</u>	<u>N</u>	<u>I</u>	<u>K</u>	<u>I</u>	<u>R</u>	<u>K</u>	<u>W</u>	<u>G</u>	<u>N</u>	<u>C</u>	<u>*</u>			
201	AGAGGACAAC	ATTAAAATTA	GAAAATGGGG	TAACTGCTGA	GGAGCCCGTA											
251	TTAAATCTTG	TAAAGCTTCA	GACGCTGAAA	TATTTTGCTA	CGTGATAATG											
301	AAGGCAATAG	AAAAATCAAT	CAGATGAAAA	ATAAACCTTG	TGCAATCCAA											
351	TGTACAGACA	TGTTACAAAA	AAGAAAAAAA	AAAAAAAAAA	AAAAAAA											

Fig. 1. (Continued)

incubated in dark at room temperature for 30 min. The alkylation reaction was quenched with DTT solution at 10 mM (final concentration). Later, the sample was diluted 10 times with 200 mM ammonium bicarbonate and digested with 0.5 μ L of trypsin solution (1 μ g/ μ L) at 37 °C for 8 h. The digestion reaction was stopped by adding 10 μ L of 2.5% TFA solution and then cleaned up with C18 ZipTip. In brief, the trypsin-digested sample was drawn to a previously wet C18 ZipTip at least 10 times. Then, the ZipTip was washed with 20 μ L of 0.1% TFA for three times. Tryptic

fragments were eluted with 20 μ L of 80% ACN/0.1% TFA by pipetting up and down for 10 times. Ten microliter of eluted fragments were subjected for LC-MS/MS analysis as described above.

2.7. Trypsin inhibitor assay

HPLC fractions (400 μ L), containing proteins with molecular masses higher than 5 kDa, were dried in a vacuum concentrator.

G) CCKP-7a

				M	K	T	L	I	S	S	V	L	L
1	GCAGCAAAAG	AATAAGAAGC	CATGAAGACT	CTGATCTCCT	CCGTCTTGCT								
	L	C	I	L	A	T	A	L	A	P	L	P	S
51	CCTCTGTATA	CTGGCAACAG	CATTGGCGCC	GCTTCCTTCT	CAACCTCAAT								
	F	K	K	V	L	K	T	L	A	E	P	N	C
101	TTTTCAAGAA	GGTTCTTAAA	ACTCTAGCTG	AGCCAAACTG	TAAGAAGTAT								
	E	G	K	K	C	D	L	N	L	N	P	V	C
151	GAGGGAAAAA	AGTGTGATCT	AAACCTAAAT	CCTGTGTGTG	GAACAAATGG								
	R	T	Y	Y	N	E	C	A	L	C	V	F	I
201	AAGAACGTAC	TACAATGAGT	GCGCGCTGTG	CGTCTTTATC	AGGGACAGTA								
	K	K	S	D	K	M	V	K	I	H	K	W	G
251	CGAAGAAATC	TGATAAGATG	GTTAAAATAC	ATAAATGGGG	TAAATGCTGA								
301	GGAGCTCAGA	ATAAATCCTG	GAGAGCTTCA	GACACCGATG	TTCTGTGAAG								
351	TGATGATTAC	AGCGAAAGAA	GAATCAATCA	GATGAAAAAT	AATCCAATGT								
401	CCAGACATGT	TCCAATAAAG	AGTGAAGCC	TTGATCTGTA	ACACAAAAAA								
451	AAAAAAAAAA	AAAAAAAAA											

Fig. 1. (Continued)

Fractions were reconstituted in 22 μ L of PBS and screened for trypsin inhibition. In brief, 180 μ L of substrate working concentration (50 μ M Z-Gly-Gly-Arg-AMC) were placed in 4 rows of a black 96-well plate. Then, 10 μ L of each fraction were placed in the plate in duplicates and the first 5 cycles of fluorescence were measured to establish a baseline. Next, 10 μ L of trypsin solution (0.001 mg/mL) were added to each well, excepting the first two that constituted negative controls and contained only substrate. The next two columns contained substrate and trypsin as positive controls. Finally, hydrolysis of the substrate by trypsin was monitored by the fluorescence generated for 60 min employing a Fluostar Optima plate reader (BMG LABTECH spectrofluorimeter) set to 460 nm emission and 395 nm excitation.

2.8. Chymotrypsin inhibition assay

The chymotrypsin inhibition assay was performed in the same way as the trypsin inhibition assay described previously with the following differences: the substrate Suc-Ala-Ala-Pro-Phe-AMC (50 μ M), and chymotrypsin solution (0.001 mg/mL) were employed for this assay.

3. Results

3.1. Molecular cloning of Kazal protein precursor encoding cDNA

A group of 7 novel Kazal-type proteins (18 variants) with proteinase inhibitory activity was identified by molecular cloning of the skin secretion of *Cruziophyla calcarifer*. Their nucleotide and translated amino acid sequences are described in Fig. 1. The translated open reading frames of these precursors contain 76 to

93 amino acids. The first 24–26 residues constitute a putative signal peptide, predicted by the SignalP 4.0 server, while the remaining 51–73 amino acids correspond to the mature Kazal-type inhibitor protein (Fig. 2). Table 1 shows the sequences of the 18 different variants found within these Kazal proteinase inhibitors and their theoretical average masses fall in the range of 5782–8392 Da. In addition, all seven Kazal-types proteins and their variants share the Kazal pattern C-X₍₇₎-C-X_(6,7)-C-X_(6,7)-Y-X₍₃₎-C-X₍₂₎-C-X_(15–21)-C, as highlighted in Table 2. Proteins were labelled *Cruziophyla calcarifer* Kazal Protein 1–7 (CCKP-1–7) adding letters for their variants. Sequences were submitted to GenBank (NCBI) with accession numbers: KX065060–KX065077.

3.2. Identification, isolation, and structural characterization of the trypsin inhibitory Kazal-type protein

The HPLC fractions of the Costa Rican population obtained in 1999 were analysed by MALDI-TOF mass spectrometry in 2013, and all fractions containing *m/z* peaks between 5 and 8 kDa were selected. The selected fractions were: 29–33. Each fraction was tested for trypsin and chymotrypsin inhibitory activity, however only fractions # 32 and 33 showed trypsin inhibitory activity (Fig. 3). MALDI-TOF mass spectrometric analysis of trypsin inhibitor fractions: #32 and 33 revealed a singly-charged ion of *m/z* 5926.43 that was confirmed by LCQ analysis (data not shown). This mass was consistent with the sequence of CCKP-1 (Table 1 and Fig. 4).

Recently, the Ecuadorian sample of *Cruziophyla calcarifer* skin secretions was fractionated by reverse phase HPLC for 240 min, and they were found to be a complex mixture of different peptides due to the numerous peaks observed in the chromatogram at 214/

	SIGNAL PEPTIDE	MATURE PEPTIDE
	1	2
CCKP-1	MKTLISSALLFCILAAALWPVPEAEG	AVSAECARYGLACNKMLAPVCGTDGTTYSNQCMLCYNNRKNKNKIEIRSRGRC*
CCKP-2	MKTLISSALLFCILAAALWPVPEAEG	ATEPDCKKYPGKCLAQNPVCGTDGRMYNECALCVFMRDSKNKVKIKKMGKC*
CCKP-3	MKTLISSVLLFCILAAALWPVPEAEG	ATPKPCPSLFSSGCPSTQDFVCGTDGNSYMNCEVMCKMKNNGGKVKVKKGYC*
CCKP-4a	MKTLISSALLFCILAAALWPVLEA	GGVVLLDCRPYGPVCSKIFDPVCGTNFITYDNTCELCKAQRENPRISMRTKGKC*
CCKP-5a	MKTLISSVLLFCILATLWPIHEAEA	VIEPNCKYEGKKCDLNPVCGTNGREYFNECALCVFIRDSKKKADKMVKIKKWKGC*
CCKP-6a	MKTLISSALLFCILAAALWPIPEVEG	EEDVACPWWYVFGCHDKYTVCGTDGVTYPNCKTLCKINGEDNIKIRKWGNC*
CCKP-7a	MKTLISSVLLLCILATALA	PLPSQPQFFKKVLKTLAEPNCKYEGKKCDLNPVCGTNGRTYYNECALCVFIRDSTKSKDKMVKIKKWKGC*

Fig. 2. Precursor structures of Kazal-type proteins from *Cruziophyla calcarifer*. All amino acid sequences contain a signal peptide (19–26 residues)-predicted by the SignalP4.1 server, followed by the mature sequence (51–73 residues).

Table 2

Alignment of the mature sequences of Kazal proteins indicating the canonical Kazal motif.

Protein	Mature peptide	I	II	III	IV	V	VI
CCKP-1 (7)	-- AVS AEC A- RYGLA C N K M L A P V C G T D G T T S N Q C M L C Y Y N R K N K K N - - - I E I R S R G R C						
CCKP-2 (8)	-- ATE PDC CKYPG- K C P L A Q N P V C G T D G R M Y Y N E C A L C V F I R D S K K K A D K M V K I K K W G K C						
CCKP-3 (7)	-- ATK PKC PSLFSSG C P S T Q D F V C G T D G N S Y M N E C V M C K M N K N N G - - - G K V K V V K K G Y C						
CCKP-4a (1)	GGVVL LDCR- PYGPV CS K I F D P V C G T N F I T Y D N T C E L C K A Q R E N P R - - - I S M R T K G K C						
CCKP-4b (1)	-- VVR LDCR- PYGPV CS K I F D P V C G T N F I T Y D N T C E L C K A Q R E N P R - - - I S M R T K G K C						
CCKP-4c (1)	-- VVR LDCR- PYGPV CS K I F D P V C G T N F I T Y D N T C E L C K A Q R E N P R - - - I S M R T K G D C R K P Y L I P E N F R R						
CCKP-5a (4)	-- V I E P N C K K Y E G K K C D L N P N P V C G T N G R E Y F N E C A L C V F I R D S K K K A D K M V K I K K W G K C						
CCKP-5b (2)	-- V I E P N C K K Y E G K K C D L N P N P V C G T N G R E Y F N E C A L C V F I R D S K K K A D K M V K I K K W G K C						
CCKP-5c (3)	-- V I E P N C K K Y E G K K C D L N P N P V C G T N G R E Y F N E C A L C V F I R D S K K K A D K M V K I K K W G K C						
CCKP-5d (1)	-- V I E P N C K K Y E G K K C D L N P N P V C G T N G R E Y F N E C A L C V F I R D S K K K A D K M V K I K K W G K C						
CCKP-6a (1)	-- E E D V A C P W Y Y V F G C H D - - K Y T V C G T D G C T Y P N K C T L C K I N G E D - - - N I K I R K W G N C						
CCKP-6a (2)	-- E E D V T C P W Y Y V F G C H D - - K Y T V C G T D G V T Y P N K C T L C K I N G E D - - - N I K I R K W G N C						
CCKP-7a (5)	P L P S Q P Q F F K K V L K T L A E P N C K K Y E G K K C D L N L N P V C G T N G R T Y Y N E C A L C V F I R D S T K K S D K M V K I H K W G K C						
CCKP-7b (1)	P P S Q P Q F S N K V L K T L A E P N C K K Y E G K K C D L N L N P V C G T N G R T Y Y N E C A L C V F I R D S T K K A D K M V K I H K W G E C						
CCKP-7c (1)	P L P S Q P Q F F K K V L K T L A E P N C K K Y E G K K C D L N L N P V C G T N G R T Y Y N E C A L C V F I R N S T K K S D K M V K I H K W G K C						
CCKP-7d (1)	P L P Q P Q F F K K V L K T L A E P N C K K Y E G K K C D L N L N P V C G T N G R T Y Y N E C A L C V F I R D S T K K S D K M V K I H K W G K C						
CCKP-7e (2)	P L P S Q P Q F F K K V L K T L A E P N C K K Y E G K K C D L N L N P V C G T N G T Y Y N E C A L C V F I R D S T K K S D K M V K I H K W G K C						
CCKP-7f (1)	P L P S Q P Q F S D K V L K T L A E P N C K K Y E G K K C D L N L N P V C G T N G R T Y Y N E C A L C V F I R D S T K K A D K M V K I H K W G K C						
Kazal pattern							
Predicted disulfide bonds							

*Conserved sites, (x) number of clones with the same sequence. The Kazal motif is highlighted: 6 Cysteines are in yellow, one Tyrosine in pink, and P1 and P2 sites in red. Connectors indicate predicted disulfide bond formation.

280 nm (Fig. 5). Besides, it was also noticeable that these peaks were very low compared to other chromatograms obtained in our research group that were 8 times higher. Usually 5–10 mg of dried secretion is employed for HPLC fractionation; however, the secretions of *C. calcarifer* were very scarce and could not be weighed. MALDI-TOF analysis was performed on each fraction and those fractions containing peptides with high molecular mass (5–8 kDa) were analysed for proteinase inhibitory activity (trypsin and chymotrypsin inhibition). The fractions tested were #81–85, 109, 110, 121–127, and 129; however, only fractions # 81–85 showed trypsin inhibition (Fig. 3). MALDI TOF MS analysis and LCQ MS confirmed a mass of 5953 Da that did not coincide with any of the sequences found by molecular cloning to date. As a consequence of the initial low amount starting material, it was

not possible to determine the K_i value of these peptidase inhibitors.

In addition, we analysed 13 skin secretion samples (1 Ecuadorian and 12 Costa Rican samples) directly by tandem mass spectrometry. By using Sequest algorithm, we confirmed the 100% coverage of amino acid sequences of the 18 variants of Kazal type proteins initially found by molecular cloning (Table 1 include the best scores of these analysis). Moreover, the fraction 91 was trypsin digested and later analysed by LC-MS/MS which confirmed the sequence of CCKP-7b (Fig. 6).

3.3. Bioinformatic analyses on Kazal-type proteinase inhibitors

The Kazal motif was identified in the 7 novel proteins and 18 variants of *Cruziophyla calcarifer* by BLAST analysis using databases from the NCBI.

The precursor sequence of *Cruziophyla calcarifer* Kazal protein-1 (CCKP-1) was compared with NCBI databases (BLAST/n tool), and it showed 76% similarity with the trypsin inhibitor, ACKT1, of *Agalychnis callidryas* (accession number HE653907.1). However, the translated mature protein only shared 52–55% similarity with trypsin inhibitor proteins from the fulmar *Fulmarus glacialis*, the falcon *Falco cherrug*, the egret *Egretta garzetta*, and *A. callidryas* (accession numbers: XP_009582735.1, XP_005436761.1, XP_009633154.1, and CCF72386.1 respectively) (Table 3). Subsequent bioinformatic analysis of the CCKP-1 predicted secondary structure showed additional similarities with the pancreatic secretory trypsin inhibitor of the salmon *Salmo salar* and the trypsin inhibitor of the boar *Sus scrofa*, especially at their C-terminal sites, as shown in Fig. 7.

The mature sequence of *Cruziophyla calcarifer* Kazal protein-2 (CCKP-2) displayed 66–68% similarity with the proteinase inhibitors: PSKP-1 and PSKP-2 of *Phyllomedusa sauvagii*; and PI01 and PI02 of *P. nordestina* (accession numbers: P83579.2, P83578.1, AFY11406.1, and AFY11405.1 respectively). In addition, CCKP-2 was 47% similar to the chymotrypsin inhibitor from the tinamou *Tinamus guttatus* (accession number KGL85187.1). Moreover, its secondary structure was similar to the chymotrypsin inhibitor of the penguin *Pygoscelis adeliae*, especially in its N-terminal region (see Fig. 8).

The mature sequence of *Cruziophyla calcarifer* Kazal protein-3 (CCKP-3) revealed 53–55% similarity with the pancreatic secretory trypsin inhibitor of the frog *Xenopus (Silurana) tropicalis*; the serine peptidase inhibitor Kazal-type 2 of the cormorant *Phalacrocorax*

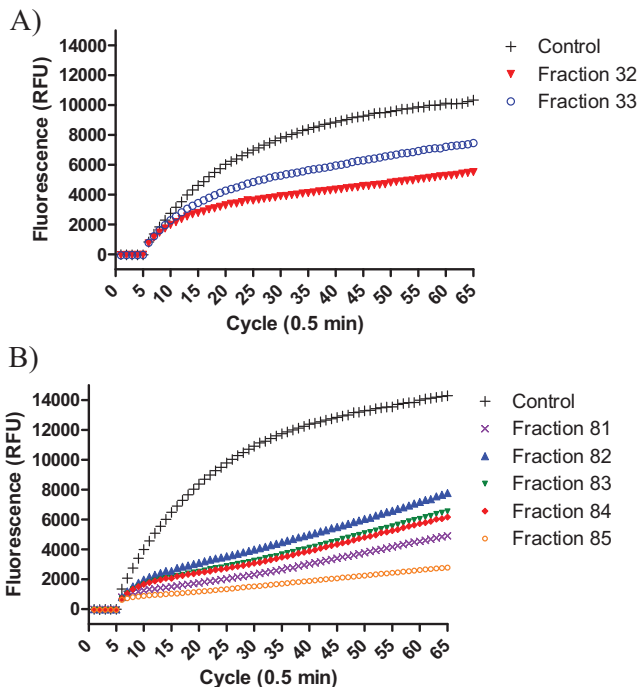


Fig. 3. Trypsin inhibitory activity of HPLC fractions. A) Fractions 32–33 from the Costa Rican sample. B) Fractions 81–85 from the Ecuadorian sample.

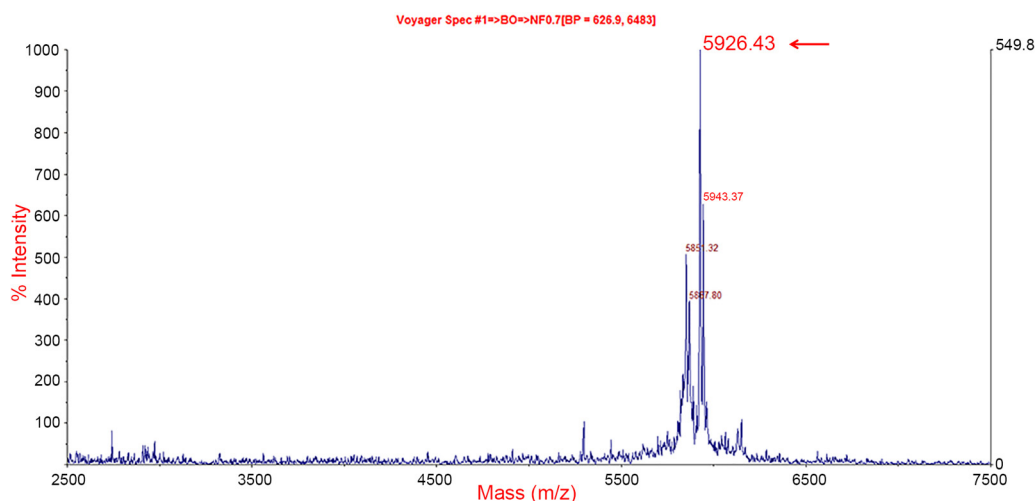


Fig. 4. MALDI-TOF MS spectrum of HPLC fraction 32 with trypsin inhibitory activity. The arrow denotes CCKP-1 with a molecular mass of 5926.43 Da.

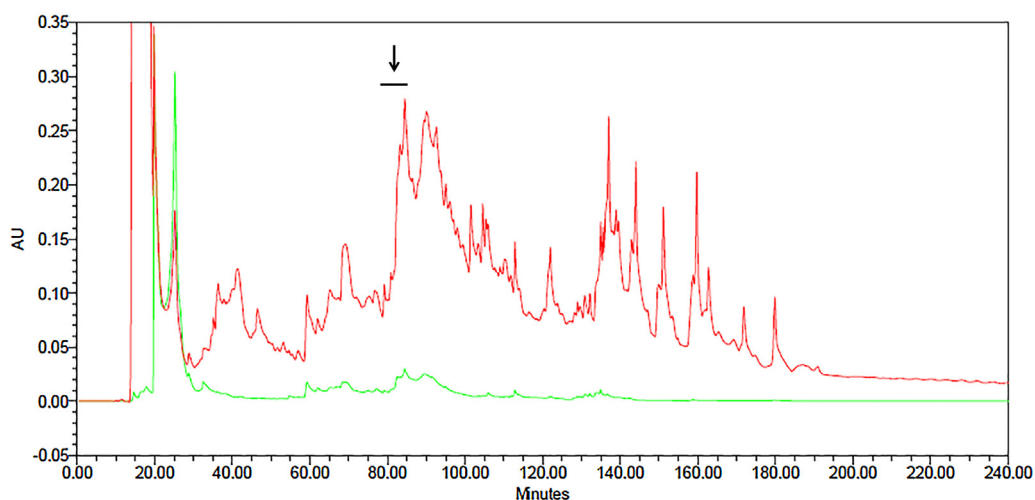


Fig. 5. Reverse phase HPLC chromatogram of skin secretion from *Cruziohyla calcarifer* fractionated over 240 min. The arrow denotes fractions 81–85 showing trypsin inhibitory activity. Detection at 214 nm (red line) and detection at 280 nm (green line).

carbo; and the trypsin inhibitor CITI-1 of the ostrich *Struthio camelus australis* (accession numbers: XP_002939857.1, XP_009500666.1, and KJV85841.1 respectively).

The mature sequences of *Cruziohyla calcarifer* Kazal protein-4 (CCKP-4a, CCKP-4b, and CCKP-4c) showed 45–52% similarity with the sperm-activating protein of the herring *Clupea pallasii*, the pancreatic secretory trypsin inhibitor of *Salmo salar*, and the serine peptidase inhibitor Kazal-type 12-like of the shrew *Tupaia chinensis* (accession numbers: BAA14009.1, NP_001140094.1, and XP_006148122.1 respectively).

The mature sequences of *Cruziohyla calcarifer* Kazal Protein-5 (CCKP-5a, CCKP-5b, CCKP-5c, and CCKP-5d) and *C. calcarifer* Kazal protein-7 (CCKP-7a to CCKP-7f), similarly to CCKP-2, exhibited 64–78% similarity with proteinase inhibitor PSKP1 and PSKP-2 of *Phyllomedusa sauvagii*; and proteinase inhibitors PI01 and PI02 of *P. nordestina* (accession numbers: P83579.2, P83578.1, AFY11406.1, and AFY11405.1 respectively).

Finally, the mature sequences of *Cruziohyla calcarifer* Kazal protein-6 (CCKP-6a and CCKP-6b) revealed 48–52% similarity with the serine peptidase inhibitor Kazal-type 2 from the cuckoo *Cuculus canorus*; the pancreatic secretory trypsin inhibitor-like from the lizard *Anolis carolinensis*; and the trypsin inhibitor CITI-1-

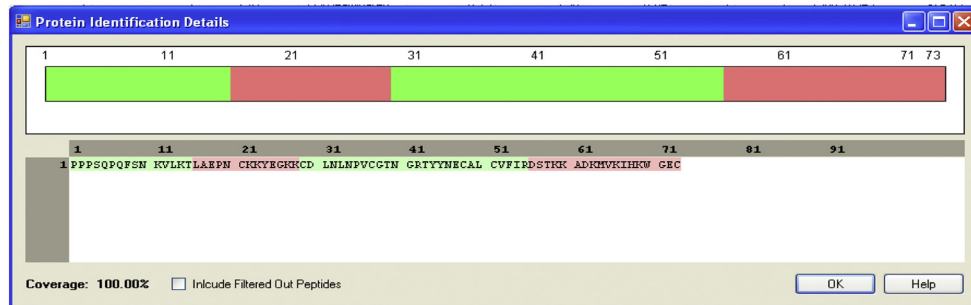
like of the bird *Cariama cristata* (accession numbers: XP_009565924.1, XP_008118717.1, and XP_009700547.1 respectively).

4. Discussion

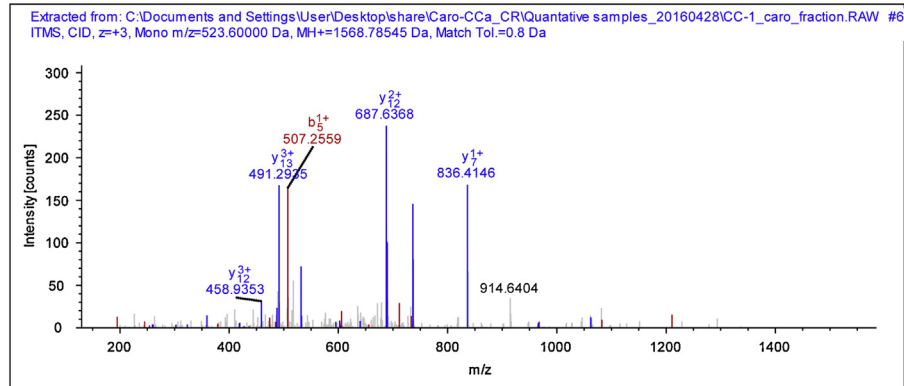
Peptidase inhibitors with three different motifs – Kunitz, Bowman-Birk and Kazal – have been previously identified in some amphibians; however, Kazal-type inhibitors appeared to be specific to Phyllomedusinae. Among them are: the Kazal prolyl endopeptidase inhibitors PSKP-1 and PSKP-2 isolated from *Phyllomedusa sauvagii*; the proteinase inhibitor PI01 and PI02 from *P. nordestina*; and the trypsin inhibitor ACKTI isolated from *Agalychnis callidryas* [18–20]. Here, we report the finding of a group of seven distinguishable Kazal-type proteinase inhibitors in *Cruziohyla calcarifer* with 18 variants (Table 1). NCBI Database interrogation using the amino acid sequence of the mature proteins identified the Kazal motif C-X₍₇₎-C-X_(6,7)-C-X_(6,7)-Y-X₍₃₎-C-X₍₂₎-C-X_(15–21)-C in all 7 Kazal-type proteins and in their 18 different variants (Table 2).

It was found that the CCKP-1 nucleotide sequence was 75% similar to the precursor of the trypsin inhibitor Kazal-type of

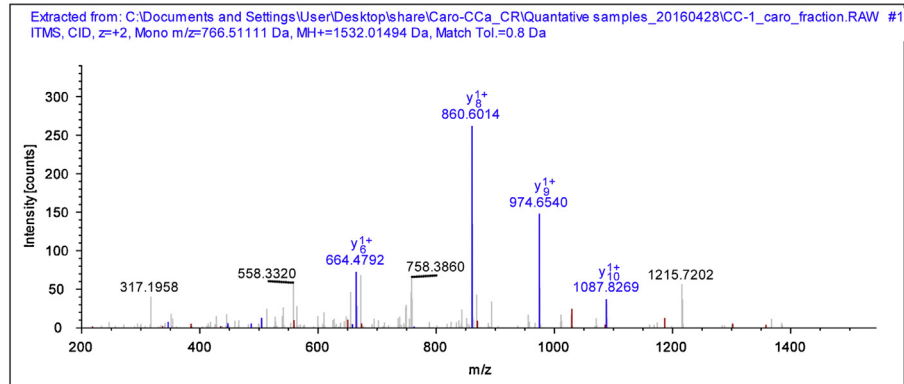
A)



B) Sequence: PPPSQPQFSNKVLK



C) Sequence: CDLNLNPVCGTNGR



D) Sequence: TYNECALCVFIR

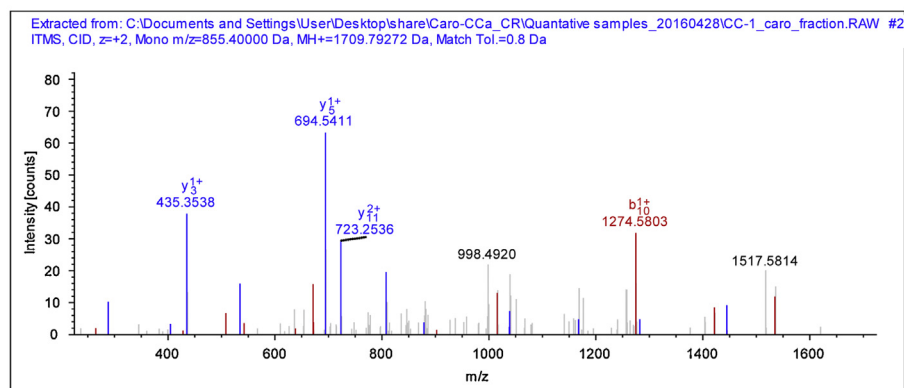


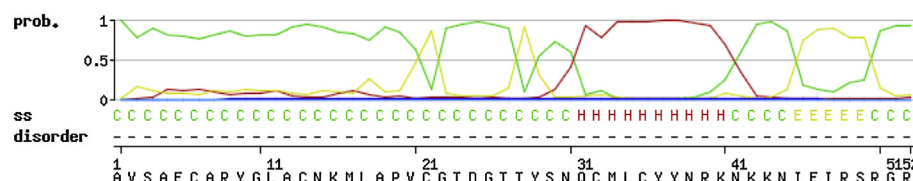
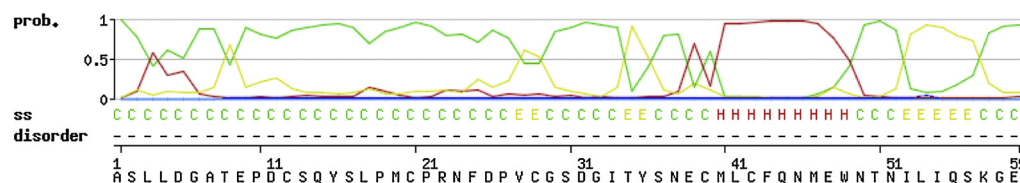
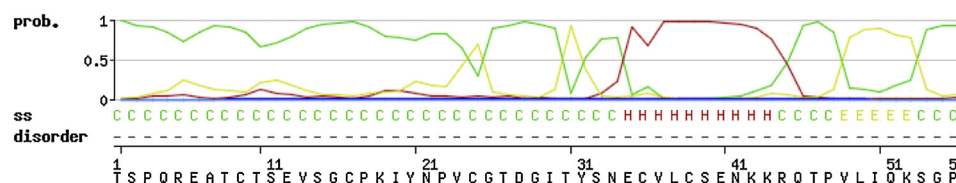
Fig. 6. Tryptic fragments sequenced by tandem mass spectrometry corresponding to the *Cruziophyla calcarifer* Kazal protein 7b (CKKP-7b). A) Sequence coverage of tryptic fragments, B–D) MS/MS profiles of the tryptic peptides sequenced by tandem mass spectrometry. Fragmented peptide ions b+ and y+ are highlighted in red and blue respectively.

Table 3

Comparison of CCKP-1 with other Kazal trypsin inhibitors according to BLAST/p (protein–protein blast).

Protein	Sequence	Identity %	Accession Number
Kazal protein <i>Cruziophyla calcarifer</i> 1 CCKP-1	A V S A E C A R Y G - - L A C N K M L A P V C G T D G T T Y S N Q C M L C Y Y N R K N K K N I E I R S R G R C		KX065060
Trypsin inhibitor precursor <i>Agalychnis callidryas</i>	A T K P R C - Q Y - - I V L C P R I L R P V C G T D G I T Y P N E C L C Q S N R D D E K D V K I Q S Q G R C	52	CCF72386.1
Pancreatic secretory trypsin inhibitor precursor <i>Salmo salar</i>	A T E P D C S Q Y S - L P M C P R N F D P V C G S D G I T Y S N E C M L C F Q N M E W N T N I L I Q S K G E C	48	NP_001140094.1
PREDICTED: trypsin inhibitor CITI-1-like partial <i>Fulmarus glacialis</i>	- - - P A C Y K Y G - V P G C P R D Y N P V C G T D G E T Y S N E C V L C L S N S E N K K D V E I F K M G R C	55	XP_009582735.1
PREDICTED: trypsin inhibitor CITI-1-like <i>Falco cherrug</i>	C P S P A C Y K Y G G V P G C P K D Y N P V C G T N G K T Y S N E C V L C F S N S E N K K N V Q I F K M G R C	53	XP_005436761.1
PREDICTED: serine protease inhibitor Kazal-type 2 <i>Egretta garzetta</i>	L Q Q P A C H R Y G - V P G C P R D Y N P V C G T D G H T Y S N E C V L C L S N S E N K K D V Q I F K M G R C	55	XP_009633154.1
Kazal pattern	C P2 P1 X X X X X C X X X X X Y X X X C X X C		

*Conserved sites, The Kazal motif is highlighted: 6 Cysteines are in yellow, one Tyrosine in pink, and P1 and P2 sites in red.

A) CCKP-1 *Cruziophyla calcarifer* Kazal Protein-1 Trypsin inhibitor**B) Pancreatic secretory trypsin inhibitor of *Salmo salar*****C) Trypsin inhibitor of *Sus scrofa***

Secondary Structure
 H helix
 E extended-beta
 C coil

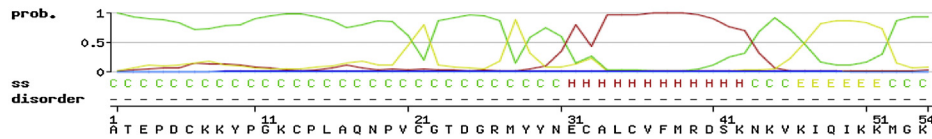
Disorder
 * disorder
 - order
 filter
 output

Fig. 7. Secondary structure prediction analysis of trypsin inhibitors using GOR IV method [22]. A) CCKP-1 *Cruziophyla calcarifer* Kazal Protein-1 Trypsin inhibitor. B) Pancreatic secretory trypsin inhibitor of *Salmo salar*. C) Trypsin inhibitor of *Sus scrofa*.

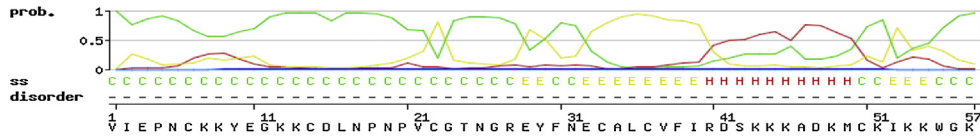
Agalychnis callidryas (ACKTI). In addition, the CCKP-1 mature sequence showed 50% similarity with a number of trypsin inhibitors from *Fulmarus glacialis*, *Falco cherrug*, *Egretta garzetta*, and *A. callidryas*. These trypsin inhibitors have a characteristic lysine (K) or arginine (R) in the P1 position and a proline (P) in the

P2 position of the active site. A trypsin inhibitory activity was predicted for CCKP-1 because it also shared a lysine (K) in the P1 position; but, in contrast, it had an asparagine (N) instead of Proline (P) in the P2 position (Table 3). Moreover, CCKP-1 had a similar secondary structure to the trypsin inhibitors of *Salmo salar*

A) CCKP-2



B) CCKP-5a



C) CCKP-7a

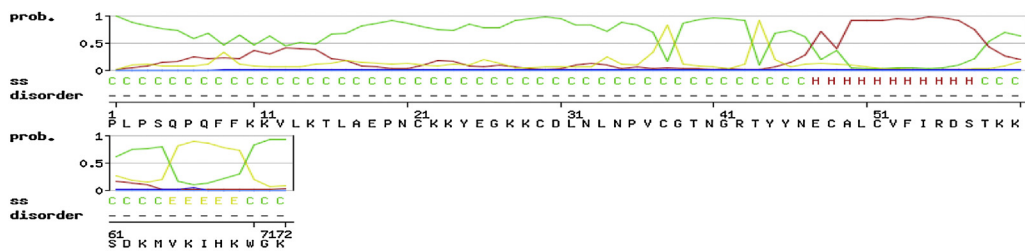
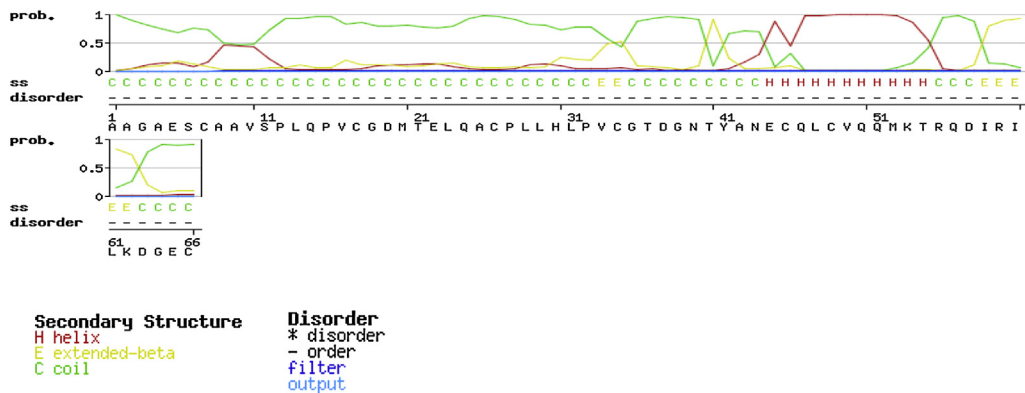
D) Chymotrypsin inhibitor of *Pygoscelis adeliae*

Fig. 8. Secondary structure prediction of chymotrypsin inhibitors using GOR IV method [22]. A) CCKP-2 *Cruziophyla calcarifer* Kazal Protein-2. B) *Cruziophyla calcarifer* Kazal Protein-5a) *Cruziophyla calcarifer* Kazal Protein-7a) Chymotrypsin inhibitor of *Pygoscelis adeliae*.

and *Sus scrofa* (Fig. 7). In addition, CCKP-1 was identified in the archived fractions 32 and 33, from an HPLC fractionation performed 15 years ago, due to their monoisotopic molecular mass of 5926.43, as determined by MALDI-TOF MS analysis and confirmed by a LCQ ESI MS (Fig. 4). Additionally, trypsin inhibitory activity was detected in those fractions confirming the activity prediction as is shown in Fig. 3.

The closest relatives of CCKP-2, based on their sequence similarities, are the proteinase inhibitors PI01 and PI02 from *Phyllomedusa nordestina*, and PSKP-1 and PSKP-2 from *P. sauvagii*. PI01 and PI02 have proline in P2 position and leucine or valine in position P1. However, PSKP-1 and PSKP-2 have two prolines in positions P1 and P2 displaying a prolyl endopeptidase activity and lacking any trypsin, chymotrypsin, V8-peptidase, and proteinase K inhibitory activity. Moreover, trypsin inhibitor activity can be

acquired if the P1 Proline is changed for lysine. In contrast, CCKP-2 has a leucine in its P1 site and a proline in the P2 site of the active site. CCKP-2 also showed 47–53% similarity with chymotrypsin inhibitors from birds, such as *Tinamus guttatus* and *Tauraco erythrophus*, which have a leucine in P1 and proline in P2 position. In addition, CCKP-2 showed a similar secondary structure to the chymotrypsin inhibitor of the penguin *Pygoscelis adeliae* (Fig. 8). Due to these sequence similarities and secondary structures, a chymotrypsin inhibitory activity is suspected for CCKP-2.

In a similar way, CCKP5 and CCKP-7 displayed a high similarity (64–81%) with the same peptidase inhibitors PI01, PI02 from *Phyllomedusa nordestina* and PSKP-1 and PSKP-2 from *P. sauvagii*. However, in contrast with CCKP-2, CCKP5 and CCKP-7 have an aspartic acid in the P2 position instead of proline, but share the

leucine in P1 position. For this reason and their similarity to the secondary structure of the chymotrypsin inhibitor of *Pygoscelis adeliae*, especially in their N-terminal regions, it is presumed that CCKP-5 and CCKP-7 also have chymotrypsin inhibitory activity similar to CCKP-2 (Fig. 8).

In the same way, CCKP-4 has a lysine (K) in position P1 which leads to a suspicion of trypsin inhibitory activity; but instead of the P2 proline, present in most other trypsin inhibitors, CCKP-4 has a serine (S). In addition, CCKP-4 showed an unusual 47% similarity to the sperm-activation protein from the herring *Clupea pallasii* most likely by sharing the Kazal motif rather than for a functional relationship.

The other two proteins, CCKP-3 and CCKP-6, have a serine (S) and an aspartic acid (D) respectively in their P1 positions. Their sequences are similar to serine peptidase inhibitors in general, but we cannot infer any specificity based on their P1 residues.

To confirm the biological activity of CCKP-2 to CCKP-7, all peptides should be identified in the individual fractions and tested. Another option could be to overexpress the 5 proteins in *Escherichia coli*, so that the purified proteins could be tested over different peptidases to confirm our predictions. Unfortunately, the scarce material available did not allow the protein identification in the individual fractions. However, their primary structures were 100% confirmed by LCQ MS/MS sequencing employing the whole secretion of 13 individual samples including wild and captive frogs from different sources. It is noticeable, that differential expression was observed in the 12 Costa Rican samples in spite of being analysed at the same concentration (1 mg/mL). Would be interesting to test in the future how the developmental stages (juvenile vs. adults), habitats, and distributions affect the expression of these protease inhibitors.

In contrast with other species, *Cruziohyala calcarifer* has proven to be a rich source of Kazal-type proteins with proven trypsin inhibitor activity in at least one of them, and predicted trypsin and chymotrypsin inhibitory activity in some others. Moreover, their extraordinary diversity contains 2 Kazal-type proteins that show unusual P1 and P2 residues which preclude their activity prediction. The biological role of those inhibitors is uncertain; however, they probably regulate the balance of antimicrobial and other bioactive peptides on the skin surface of the frogs, to protect them against pathogenic microorganisms and predators. The extraordinary diversity of these Kazal-type protein inhibitors in the skin of *C. calcarifer* possibly mirrors the complexity of peptides and proteins still unknown in other species, and promotes research on this species in particular.

5. Conclusions

In conclusion, the combined strategy including tandem mass sequencing, molecular cloning and HPLC fragmentation allowed the identification of 18 variants of Kazal-type proteins of *Cruziohyala calcarifer* (CCKP). These proteins were classified in 7 types by sequence similarity. CCKP-1 and 4 were deemed of trypsin inhibitory activity while CCKP-2, 5, 7 were deemed of chymotrypsin inhibitory activity by sequence homology with other peptidase inhibitors. In addition, CCKP-1 was identified in the chromatographic fraction No.32 showing trypsin inhibitory activity and a molecular mass of $[M+H]^+ = 5926.43$ Da. In this way, *C. calcarifer* has showed to be an important source of diverse peptidase inhibitors promoting further research of the skin secretions in this species.

Conflict of interest

The authors declare that there is no conflict of interest.

Authorship

This study was conceived and designed by CS, MZ, TC. Sample collections were performed by CPB, EET, and LAC. Data were acquired by CPB and RL. LC-MS/MS analysis was performed by LW and XX. The article was written by CPB and reviewed critically by CS and LAC.

Acknowledgments

Carolina Proaño-Bolaños is in receipt of a scholarship of the Ecuadorian Secretariat of Science and Technology (SENESCYT). This research was funded by the Natural drug discovery group, School of Pharmacy, Queen's University Belfast and SENESCYT. The latter also supported field work to CPB. Collection and rearing of frogs, transportation and analysis of samples were done under permits of the Ecuadorian Ministerio de Ambiente (MAE): 001-13 IC-FAU-DNB/MA,003-11-IC-FAU-DNB/MA, 005-15 IC-FAU-DNB/MA.; Framework contract for access to genetic resources MAE-DNB-CM-2016-0051; and exportation permits: 003-13-EXP-CI-FAU-DNB/MA and 2015-003-FO-DPAP-MA. This research is part of the MAE project "Conservation of Ecuadorian amphibian diversity and sustainable use of its genetic resources", which is promoted by The Global Environmental Facility (GEF) and Programa de las Naciones Unidas para el Desarrollo (PNUD), and that involves Ikiam-Universidad Regional Amazónica, Queen's University Belfast, and Centro Jambatu.

Thanks to Mayra Rojas and David Narvaez of Universidad de las Américas in Quito, Ecuador, who helped to prepare freeze-dried samples. Christian Proy (Austria), and Stassen Raf (Belgium) generously provided access to their Costa Rican pet frogs to obtain samples for this research. Thanks to Dr. Catriona Arlow for reviewing the language of the first draft of this manuscript.

References

- [1] V. Ersparmer, Bioactive secretions of the amphibian integument, in: H. Heatwole, G. Bartholameus (Eds.), *Amphibian Biology. The Integument*, vol. 1, Surrey, Beatty and Sons, Chipping-Norton, N.S.W, 1994, pp. 178–350.
- [2] A.H. Thompson, A.J. Bjorson, D.F. Orr, C. Shaw, S. McClean, A combined mass spectrometric and cDNA sequencing approach to the isolation and characterization of novel antimicrobial peptides from the skin secretions of *Phyllomedusa hypochondrialis azurea*, *Peptides* 28 (2007) 1331–1343.
- [3] X. Chen, L. Wang, H. Wang, H. Chen, M. Zhou, T. Chen, C. Shaw, A fish bradykinin (Arg0, Trp5, Leu8-bradykinin) from the defensive skin secretion of the European edible frog, *Pelophylax kl. esculentus*: structural characterization; molecular cloning of skin kininogen cDNA and pharmacological effects on mammalian smooth muscle, *Peptides* 32 (2011) 26–30.
- [4] M.A. Apponyi, T.L. Pukala, C.S. Brinkworth, V.M. Maselli, J.H. Bowie, M.J. Tyler, G.W. Booker, J.C. Wallace, J.A. Carver, F. Separovic, J. Doyle, L.E. Llewellyn, Host-defence peptides of Australian anurans: structure, mechanism of action and evolutionary significance, *Peptides* 25 (2004) 1035–1054.
- [5] M. Laskowski Jr., I. Kato, Protein inhibitors of proteinases, *Annu. Rev. Biochem.* 49 (1980) 593–626.
- [6] C.J. Farady, C.S. Craik, Mechanisms of macromolecular protease inhibitors, *ChemBiochem* 11 (2010) 2341–2346.
- [7] W. Bode, R. Huber, Natural protein proteinase inhibitors and their interaction with proteinases, *Eur. J. Biochem.* 204 (1992) 433–451.
- [8] R. Lai, H. Liu, W.H. Lee, Y. Zhang, Identification and cloning of a trypsin inhibitor from skin secretions of Chinese red-belly toad *Bombina maxima*, *Comp. Biochem. Physiol. B: Biochem. Mol. Biol.* 131 (2002) 47–53.
- [9] T. Chen, C. Shaw, Identification and molecular cloning of novel trypsin inhibitor analogs from the dermal venom of the Oriental fire-bellied toad (*Bombina orientalis*) and the European yellow-bellied toad (*Bombina variegata*), *Peptides* 24 (2003) 873–880.
- [10] M. Liberio, I. Bastos, J.O. Pires, W. Fontes, J.M. Santana, M.S. Castro, The crude skin secretion of the pepper frog *Leptodactylus labyrinthicus* is rich in metallo and serine peptidases, *PLoS One* 9 (2014) e96893.
- [11] G. Demesa-Balderama, E.P. Meneses, L. Hernandez-Orihuela, V. Pando-Robles, M.C. Rodriguez, C. Barrientos-Salcedo, M.B. Aguilar, C.V. Batista, A comprehensive proteomic study of the skin secretions of the frog *Lithobates spectabilis*, *Protein Pept. Lett.* 23 (2016) 597–611.
- [12] N.J. Darby, D.B. Lackey, D.G. Smyth, Purification of a cysteine endopeptidase which is secreted with bioactive peptides from the epidermal glands of *Xenopus laevis*, *Eur. J. Biochem.* 195 (1991) 65–70.

- [13] G. Mignogna, S. Pascarella, C. Wechselberger, C. Hinterleitner, C. Mollay, G. Amiconi, D. Barra, G. Kreil, BSTI, a trypsin inhibitor from skin secretions of *Bombina bombina* related to protease inhibitors of nematodes, *Protein Sci.* 5 (1996) 357–362.
- [14] X. Lu, Y. Ma, J. Wu, R. Lai, Two serine protease inhibitors from the skin secretions of the toad, *Bombina microdeladigitata*, *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 149 (2008) 608–612.
- [15] M.F. Ali, K.R. Lips, F.C. Knoop, B. Fritzsche, C. Miller, J.M. Conlon, Antimicrobial peptides and protease inhibitors in the skin secretions of the crawfish frog, *Rana areolata*, *Biochim. Biophys. Acta (BBA)—Proteins Proteom.* 1601 (2002) 55–63.
- [16] J.M. Conlon, J.B. Kim, A protease inhibitor of the Kunitz family from skin secretions of the tomato frog, *Dyscophus guineti* (Microhylidae), *Biochem. Biophys. Res. Commun.* 279 (2000) 961–964.
- [17] H. Wang, L. Wang, M. Zhou, M. Yang, C. Ma, T. Chen, Y. Zhang, M. Zeller, M. Hornshaw, C. Shaw, Functional peptidomics of amphibian skin secretion: a novel Kunitz-type chymotrypsin inhibitor from the African hyperoliid frog, *Kassina senegalensis*, *Biochimie* 94 (2012) 891–899.
- [18] G. Song, M. Zhou, W. Chen, T. Chen, B. Walker, C. Shaw, HV-BBI-a novel amphibian skin Bowman-Birk-like trypsin inhibitor, *Biochem. Biophys. Res. Commun.* 372 (2008) 191–196.
- [19] M. Wang, L. Wang, T. Chen, B. Walker, M. Zhou, D. Sui, J.M. Conlon, C. Shaw, Identification and molecular cloning of a novel amphibian Bowman Birk-type trypsin inhibitor from the skin of the Hejiang Odorous Frog; *Odorrana hejiangensis*, *Peptides* 33 (2012) 245–250.
- [20] J. Li, J. Wu, Y. Wang, X. Xu, T. Liu, R. Lai, H. Zhu, A small trypsin inhibitor from the frog of *Odorrana grahami*, *Biochimie* 90 (2008) 1356–1361.
- [21] L.G. Gebhard, F.U. Carrizo, A.L. Stern, N.I. Burgardt, J. Faivovich, E. Lavilla, M.R. Ermacora, A Kazal prolyl endopeptidase inhibitor isolated from the skin of *Phyllomedusa sauvagii*, *Eur. J. Biochem.* 271 (2004) 2117–2126.
- [22] R. Li, H. Wang, Y. Jiang, Y. Yu, L. Wang, M. Zhou, Y. Zhang, T. Chen, C. Shaw, A novel Kazal-type trypsin inhibitor from the skin secretion of the Central American red-eyed leaf frog, *Agalychnis callidryas*, *Biochimie* 94 (2012) 1376–1381.
- [23] M. Neiva, D.C. Vargas, K. Conceicao, G. Radis-Baptista, M.T. Assakura, C. Jared, M.A. Hayashi, Gene expression analysis by ESTs sequencing of the Brazilian frog *Phyllomedusa nordestina* skin glands, *Toxicon* 61 (2013) 139–150.
- [24] Y.H. Abdel-Wahab, L. Marenah, D.F. Orr, C. Shaw, P.R. Flatt, Isolation and structural characterisation of a novel 13-amino acid insulin-releasing peptide from the skin secretion of *Agalychnis calcarifer*, *Biol. Chem.* 386 (2005) 581–587.
- [25] C. Proaño-Bolaños, M. Zhou, L. Wang, L.A. Coloma, T. Chen, C. Shaw, Peptidomic approach identifies cruzioseptins, a new family of potent antimicrobial peptides in the splendid leaf frog, *Cruziohyla calcarifer*, *J. Proteom.* 146 (2016) 1–13.
- [26] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, *Mol. Biol. Evol.* 30 (2013) 2725–2729.
- [27] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, *J. Mol. Biol.* 215 (1990) 403–410.
- [28] J. Rozenski, *Peptide Mass Calculator v3.2* 1999.
- [29] C. Combet, C. Blanchet, C. Geourjon, G. Deleage, NPS@ network protein sequence analysis, *TIBB* 25 (2000) 147–150.